

世界におけるAI規制の動向

藍澤 志津 ●一般財団法人マルチメディア振興センター ICTリサーチ&コンサルティング部 リサーチディレクター

開発競争力を維持しながら責任あるAIの実現を目指して米国、中国、欧州などで法制化の動きが進む。G7の広島AIプロセスをはじめ、生成AIのリスク低減に向けた国際協力の枠組みも始まっている。

2016年ごろから、各国政府・国際機関は人工知能（AI）技術の開発の進展を踏まえ、社会・経済構造の変化を想定し、AIに関する国家戦略・原則を打ち出した¹。AI原則に関しては国際的なコンセンサスが形成されつつあり、日本も統合イノベーション戦略推進会議決定において「人間中心のAI社会原則」を発表した。また、日本を含む複数国間で合意したOECDのAI原則では、①包摂的な成長、持続可能な開発および幸福②人間中心の価値観および公平性③透明性および説明可能性④堅牢性、セキュリティおよび安全性⑤アカウントビリティ——が掲げられた²。

AIの社会実装が進展する中、2022年11月、米OpenAIが、大規模言語モデル（LLM）に基づく生成AIの一種であるAIチャットボット「ChatGPT」のプロトタイプを一般公開した。これは世界で急速に利用が拡大し、今もなお人々の生活や仕事に大きな影響を及ぼしている。これを受け各国政府は、AIを構成要素として含むAIシステム、AIシステムの機能を提供するAIサービス、その他付随的サービスと、これらを開発・利用・提供する者に対する規制・ガバナンスに関して、次々に方針・対策を打ち出している。

以下に米国、中国、EUと、国際協力枠組みの動向を報告する。

■主要国におけるAI規制

●米国

【AI権利章典】

ジョー・バイデン米大統領は2021年1月20日の就任初日、連邦政府全体に対して、不公平を根絶し、意思決定プロセスに公正さを組み込み、米国における（プライバシー権を含む）公民権、機会均等、人種的公正を積極的に推進するよう命じた³。これを踏まえ、ホワイトハウスの科学技術政策局（OSTP）は2022年10月、AIの開発等における原則をまとめた「AI権利章典の青写真——アルゴリズム時代における公民権保護のビジョン」（以下、AI権利章典）を発表した⁴。AI権利章典は「技術、データ、自動化システムの利用がバイデン政権の礎石としている米国の基本原則である公民権や民主主義的価値観と引き換えにもたらされるものであってはならない」としている。その上で、AIの時代において米国民を保護するための自動化システムの設計、使用、展開の指針となる5つの原則、すなわち①安全で効果的なシステム②アルゴリズムによる差別からの保護③データのプライバシー④利用者への通知と説明⑤人間的な代替案、配慮、撤回（フォールバック）——を示した。

AI権利章典は、AIの説明責任を拡大する内容

となっている。特筆すべきは、アルゴリズムによる差別・バイアスからの保護を掲げた点である。これに先立つ2022年5月には、司法省と雇用機会均等委員会（EEOC）が、AIが生み出す雇用差別に対処するガイダンスを発表し、採用候補者の事前審査に使われるソフトウェア、アルゴリズム、AIが障害を持つ候補者を不当に差別した場合、連邦公民権法違反となる恐れがあると警告した⁵。

このAI権利章典は、実務レベルでのAIによるバイアス問題の解決への道筋を示し、連邦政府による公民権の保護の取り組みを支える内容となった。

【大手IT企業による自主的取り組みで合意】

生成AIの開発競争が加速する中、バイデン政権は2023年7月、AI技術の安全、安心、透明性のある開発への移行を支援するため、AIを開発する主要企業7社（アマゾン・ドット・コム、アンソロピック（Anthropic）、グーグル、インフレクションAI（Inflection AI）、メタ、マイクロソフト、OpenAI）による自主的な取り組みの実施で合意を取り付けたと発表した⁶。7社は、3つの原則（安全性、セキュリティ、信頼性）に基づきAI普及に伴う偽情報やその他のリスクに対処するため、AI生成コンテンツを識別できる透かしシステムを開発する等の安全策の導入を約束した。

【AIに関する包括的な大統領令】

バイデン大統領は2023年10月、AIに関する包括的な大統領令に署名した⁷。これは①AIの安全性とセキュリティを高める新基準②米国民のプライバシー保護③平等と人権の尊重④消費者、患者、学生のために立ち上がる⑤労働者の支援⑥イノベーションと競争の推進⑦米国の国際リーダーシップ向上⑧政府による責任ある効果的なAIの活用促進——という8つの原則から成っている。

同大統領令は、米国で初めて法的拘束力を持つAI規制となる。これに加え、議会においてAI規制に関する法案の策定も進められている⁸。

●中国

【責任あるAIの発展】

中国政府は2019年3月、AIに関する法律、倫理、社会問題の研究を強化し、AIの世界的なガバナンスを積極的に推進するため「国家次世代AIガバナンス専門委員会」を設立した。同委員会は同年6月には「新世代AIガバナンス原則——責任あるAIの発展」を発表し①調和（和諧）・友好②公平・公正③包摂・共有④プライバシーの尊重⑤セキュリティ・制御可能性⑥責任の共有⑦開放・協力⑧アジャイルガバナンス——という8つの原則を提示した⁹。同原則はAIガバナンスの枠組みと行動ガイドラインであり「責任あるAIの発展」を強調している。

【アルゴリズム規制】

中国政府は2021年11月、「インターネット情報サービスに関するアルゴリズム推奨の管理に関する規定」を発表した¹⁰。レコメンデーション・アルゴリズム・サービスに関して、サービス提供者に対しレコメンデーション・アルゴリズムの利用を分かりやすく明示することと、アルゴリズムの基本原則、目的、意図、構造等を適切な表現で告知することを義務付けた。

さらに、政府は2022年12月、「インターネット情報サービス深度合成管理規定」を制定し、深度合成技術¹¹に対する規則を制定した¹²。生成・編集される情報内容に関して、合理的な場所や分かりやすいサインにより、深度合成アルゴリズムの利用を明示することが義務付けられた。

【AI倫理規制】

中国政府は2022年4月、「科学技術倫理ガバナンス強化に関する意見」を発表し、倫理5原則（人間の幸福度の向上、生存権の尊重、公正・公正の堅持、リスクの合理的な管理、オープン性と透明性）を示した¹³。さらに、2023年4月には同倫理原則を立法化するべく「科技倫理審査弁法（試行）」に関する諮問が開始された¹⁴。AIを含む科学技術の倫理審査・監督の強化と、責任あるイノベーションの推進を目的とする。

【生成AI規制】

中国政府は2023年7月、ChatGPTや百度（バイドゥ）の「文心一言（ERNIE Bot）」等の生成AIの中国でのサービス提供に当たり、守るべき点や罰則を定めた「生成型AIサービス管理暫定弁法」を発表し、同年8月15日から施行した¹⁵。

同弁法での生成AIは「アルゴリズム、モデル、ルールに基づきテキスト、画像、音声、動画、コード、その他のコンテンツを生成する技術」と定義されている。

サービス提供者に義務付けられるのは、知的財産権を侵害しないこと、個人情報保護の責任を負うこと、サービス提供前にセキュリティ評価を報告しアルゴリズムを提出すること、事前学習データと最適化された学習データの合法性に責任を負うこと、利用者に身元情報の提供を義務付けること、利用者が生成コンテンツに過度に依存・耽溺しないように適切な措置を講じること、人種・国籍・性別等に基づく差別的なコンテンツを生成しないこと、生成コンテンツをラベリングすること、苦情受け付けの仕組みを構築することなどとなっている。また、中国の領土外から本土への生成AIサービスの提供が法律・行政規則等に準拠していない場合、政府は関連機関に通知し、技術的措置およびその他の必要な措置を講じる。

同弁法に違反した場合は「サイバーセキュリ

ティ法」「データセキュリティ法」「個人情報保護法」等に基づいて罰則が科される。規定がない場合には警告や是正要請が発出されるが、是正拒否や状況が深刻な場合には、サービスの停止や終了が命じられるとともに、行政処分や、刑事責任が追及される場合もある。

●欧州連合（EU）

【欧州連合AI法案】

欧州委員会は2021年4月、「AIに関する調和された規則を定める規則案」（欧州連合AI法案）を発表した¹⁶。欧州のAIを信頼に足るものにするるとともに、AIの優れたエコシステムを構築し、EUのグローバル競争力を強化することを目的としている。同法案は、AIのリスクのレベルを以下の4つに分類・定義し、それに応じてプロバイダーやAIシステムを導入する者の義務を定めるというリスクベースのアプローチを採っている¹⁷。

（1）許容できないリスク

政府による社会的採点から、危険な行動を助長する音声アシスト玩具等、人々の安全・生活・権利を明らかに脅かすAIシステム。全て禁止される。

（2）高リスク

交通機関のような重要インフラのAIシステム等。市場に投入される前に、適切なりスク評価と軽減システム、リスク最小化のための人的監視手段等の厳しい義務が課される。

（3）限定的なりスク

特定の透明性義務を有するAIシステム。例えばチャットボットを利用する場合、利用者はAIシステムの利用を認識し十分な情報を得た上で、継続か後退かを決定できるようにすべきである。

（4）最小リスクまたはリスク無し

自由な利用を認める。AI対応のビデオゲームやスパムフィルター等。現在EUで使用されている

AIシステムの大半が属する。

さらに、同法案ではAIシステムが市場投入された場合、当局が市場監視を担当すること、利用者は人的監視とモニタリングを確保すること、プロバイダーはモニタリング体制を整備すること、プロバイダーと利用者が重大な事故や故障を報告することを規定している。

同法案が発表された後、2022年11月にOpenAIがChatGPTのプロトタイプを発表し、世界的に普及したことを受け、欧州理事会は同年12月に、同法案に「汎用AI」の内容を追加した。さらに、2023年6月に欧州議会が修正案を採択し「基盤モデル」「生成AI」の内容を追加した。

その後、2023年12月、欧州議会と理事会は、同法案に関して暫定合意に達した¹⁸。主な合意内容は以下の通りである。

・禁止されるAIシステム

市民の権利と民主主義にもたらす潜在的脅威があるとして、以下のようなAIシステムを禁止する。センシティブな特性（政治的、宗教的、哲学的信条、性的指向、人種等）を利用した生体認証分類システム、顔認識データベース作成のためのインターネットやCCTV映像からの顔画像の無作為なスクレイピング、職場や教育機関における感情認識、社会的行動や個人的特徴に基づく社会的採点、人間の行動を操作して自由意志を回避するAIシステム、人の脆弱性（年齢、障害、社会的または経済的状況）を利用するAIシステム。

一方、遠隔生体認証（RBI）システムに関して、事後のRBIは重犯罪訴追のため司法の許可を得た場合に限定され、リアルタイムのRBIは被害者（誘拐、人身売買、性的搾取）の捜索、特定かつ現在のテロ脅威の防止、特定の犯罪（テロ、人身売買、性的搾取、殺人、誘拐、強姦、武装強盗、

犯罪組織への参加、環境犯罪等）を犯したと疑われる人物の位置特定や身元確認に限定して認められる。

・高リスクのAIシステム

健康、安全、基本的権利、環境、民主主義、法の支配に重大な害を及ぼす可能性があるとして高リスクAIシステムの義務を明確化した。保険や銀行部門に対して基本的人権の影響評価を義務化するとともに、選挙結果や有権者の行動に影響を与えるAIシステムも高リスクに分類し、市民は苦情の申し立てや説明を受ける権利を持つ。

・汎用AIシステムのガードレール

汎用AIシステムおよびモデルは、透明性要件に準拠する。同要件には、コンテンツがAIによって生成されたことを開示することや、技術文書の作成、EU著作権法の順守、トレーニングに使用されたコンテンツに関する詳細な要約の配布等が含まれる。影響力の大きい汎用AIモデルについては、モデル評価の実施、リスクの評価と軽減、敵対的テストの実施、重大インシデントの報告、サイバーセキュリティの確保、エネルギー効率の報告等を義務付ける。

・制裁と発効

不順守の場合、違反内容と企業規模に応じて3500万ユーロまたは世界売上高の7%から750万ユーロまたは売上高の1.5%までの罰金が科せられる可能性がある。

同法案は今後、議会と理事会の双方での正式な承認を経て成立する。なお、同法案は「規則（Regulation）」であり、加盟国の国内法制化を経ることなく直接適用されることとなる。早ければ2026年にも全面適用される予定だが、AIのリス

クと活用規制を詳細に記述しており、世界への影響力は大きいと考えられる。

■国際協力枠組み

●広島AIプロセス

2023年5月に広島で主要7か国首脳会議（G7）が開催され、利用や開発が急拡大するChatGPT等の生成AIをめぐる著作権の保護や偽情報への対応などについて閣僚級で議論する「広島AIプロセス」を開始することが盛り込まれた。これを踏まえ、同年10月にAIの開発者が守るべき責務をまとめた国際指針と行動規範で合意した¹⁹。

市場投入前から利用までの各段階でリスクを低減することを求める内容となっており、具体的には、開発企業に対して、レッドチームによる疑似的なテスト、電子透かしやAIコンテンツ識別等の技術の導入、高度なAIシステムの能力領域の公表やインシデントの報告、権利侵害を防ぐためのプライバシーや知的財産を尊重する安全措置等を求めた。

●AI安全サミット

2023年11月、英国主催の「AI安全サミット」が開催され、開催地にちなみ「ブレッチリー宣言」が発表された²⁰。日本、中国、米国を含む29か国が署名し、フロンティアAI²¹の安全かつ責任ある開発、機会とリスク、最も重要な課題に対応するための国際的行動の必要性について、世界初の合意に達した。これは、G7、OECD、欧州評議会、国連、GPAI（AIに関するグローバルパートナーシップ）を含む既存の国際的な取り組みを補完するものである。

さらに、各国政府とAI企業はフロンティアAI

モデルの安全性テスト計画で合意した。モデルの配備前と配備後にAIの安全性を確保し、特に国家安全保障、安全、社会的危害等のテストで協力する。これにより、フロンティアAIモデルの安全性に関する責任を、企業のみではなく政府も担うこととなった。

フロンティアAIの安全性に関する国際協力プロセスとして、韓国で6か月以内にAIに関するミニ・バーチャル・サミットが共同開催されることと、フランスが1年後の次回AI安全サミットのホスト国となることが決定した。

今後も、AIリスクに取り組む国際協力は推進されていく。

■責任あるAI開発の推進に向けて

以上述べてきたように、AI規制に関して、中国では世界に先駆けて生成AIに関する規制が導入されており、米国と欧州では法制化が進行中である。一方、G7の広島AIプロセスでは開発企業向けの国際指針と行動規範が合意され、英国主催のAI安全サミットでは官民協力によるAIの安全性確保が推進される等、国際協力によるAIのリスクの低減が目指されることとなった。日本でも、AI事業者向けの指針の議論が進んでおり、政府のAI戦略会議では、生成AIのリスク低減に向けて法規制も含めた検討が進んでいる。

進化し続けるAI環境において、倫理と開発のバランスを取りつつ、責任あるAI開発を推進する緊急性はかつてないほど高まっている。各国政府、企業、学界、国際社会、あらゆる利害関係者がさらなる連携を通じて、リスクを軽減しすべての人々に役立つAIの未来を確保することが望まれている。

1. マルチメディア振興センター、「世界各国におけるAIの国家戦略」、『インターネット白書2021』、2021年
2. OECD, Recommendation of the Council on Artificial Intelligence, May 22, 2019
<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
(以下、ウェブページの参照はすべて2023年12月12日)
3. The White House, Inaugural Address by President Joseph R. Biden, Jr., Jan. 20, 2021
<https://www.whitehouse.gov/briefing-room/speeches-remarks/2021/01/20/inaugural-address-by-president-joseph-r-biden-jr/>
4. The White House, Blueprint for an AI Bill of Rights: A Vision for Protecting Our Civil Rights in the Algorithmic Age, Oct. 4, 2022
<https://www.whitehouse.gov/ostp/news-updates/2022/10/04/blueprint-for-an-ai-bill-of-rights-a-vision-for-protecting-our-civil-rights-in-the-algorithmic-age/>
5. U.S. Department of Justice, Justice Department and EEOC Warn Against Disability Discrimination, May 12, 2022
<https://www.justice.gov/opa/pr/justice-department-and-eecoc-warn-against-disability-discrimination>
6. The White House, FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI, Jul. 21, 2023
<https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/>
7. The White House, Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, Oct. 30, 2023
<https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>
The White House, FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence, Oct. 30, 2023
<https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/>
8. The White House, WHAT THEY ARE SAYING: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence, Oct. 31, 2023
<https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/31/what-they-are-saying-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/>
9. 科学技术部, 发展负责任的人工智能: 新一代人工智能治理原则发布, Jun. 17, 2019
https://www.most.gov.cn/kjbgz/201906/t20190617_147107.html
10. 国家互联网信息办公室他, 互联网信息服务算法推荐管理规定, Jan. 4, 2022
http://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm
11. ディープラーニング、バーチャルリアリティ、その他の生成成型アルゴリズムを使用してテキスト、画像、音声、映像、バーチャルシーン、その他のネットワーク情報を生成する技術。
12. 国家互联网信息办公室他, 互联网信息服务深度合成管理规定, Dec. 11, 2022
http://www.cac.gov.cn/2022-12/11/c_1672221949354811.htm
13. 科学技术部, 中共中央办公厅国务院办公厅印发《关于加强科技伦理治理的意见》, Mar. 20, 2022
https://www.most.gov.cn/xxgk/xinxifenlei/fdzdgnr/fgzc/gfxwj/gfxwj2022/202203/t20220321_179899.html
14. 科学技术部, 关于公开征求对《科技伦理审查办法(试行)》意见的公告(已结束), Apr. 4, 2023
https://www.most.gov.cn/wsdc/202304/t20230404_185388.html?_ga=2.121368968.1251830295.1701752847-11659524.1696147575&_fsi=TFzHINRc
15. 国家互联网信息办公室他, 生成式人工智能服务管理暂行办法, Jul. 13, 2023
http://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm
16. European Commission, Proposal for a Regulation laying down harmonised rules on artificial intelligence, Apr. 21, 2021
<https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence>
17. European Commission, Regulatory framework proposal on artificial intelligence, n.d.
<https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
18. European Parliament, Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI. Dec. 9, 2023
<https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>
19. 外務省、広島AIプロセスに関するG7首脳声明、2023年10月30日
https://www.mofa.go.jp/mofaj/ecm/ec/page5_000483.html
20. GOV.UK., Countries agree to safe and responsible development of frontier AI in landmark Bletchley Declaration, Nov. 1, 2023
<https://www.gov.uk/government/news/countries-agree-to-safe-and-responsible-development-of-frontier-ai-in-landmark-bletchley-declaration>
21. 英国政府は「多様なタスクをこなし、現在の最先端モデルと同等かそれ以上の能力を持つ、非常に高性能な汎用AIモデル」と定義している。



1996, 1997, 1998, 1999, 2000...

[インターネット白書 ARCHIVES] ご利用上の注意

このファイルは、株式会社インプレスR&Dおよび株式会社インプレスが1996年～2024年までに発行したインターネットの年鑑『インターネット白書』の誌面をPDF化し、「インターネット白書 ARCHIVES」として以下のウェブサイトで公開しているものです。

<https://IWParcives.jp/>

このファイルをご利用いただくにあたり、下記の注意事項を必ずお読みください。

- 記載されている内容(技術解説、データ、URL、名称など)は発行当時のものです。
- 収録されている内容は著作権法上の保護を受けています。著作権はそれぞれの記事の著作者(執筆者、写真・図の作成者、編集部など)が保持しています。
- 著作者から許諾が得られなかった著作物は掲載されていない場合があります。
- このファイルの内容を改変したり、商用目的として再利用したりすることはできません。あくまで個人や企業の非商用利用での閲覧、複製、送信に限られます。
- 収録されている内容を何らかの媒体に引用としてご利用される際は、出典として媒体名および年号、該当ページ番号、発行元などの情報をご明記ください。
- オリジナルの発行時点では、株式会社インプレスR&Dおよび株式会社インプレスと著作権者は内容が正確なものであるように最大限に努めましたが、すべての情報が完全に正確であることは保証できません。このファイルの内容に起因する直接のおよび間接的な損害に対して、一切の責任を負いません。お客様個人の責任においてご利用ください。

お問い合わせ先

インプレス・サステナブルラボ

✉ iwp-info@impress.co.jp